A REVIEW AND SOME EXTENSIONS OF TAKEMURA'S GENERALIZATIONS OF COCHRAN'S THEOREM

TECHNICAL REPORT NO. 2

GEORGE P.J. STYAN

SEPTEMBER 1982

U. S. ARMY RESEARCH OFFICE
CONTRACT DAAG29-82-K-0156

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

A REVIEW AND SOME EXTENSIONS OF TAKEMURA'S GENERALIZATIONS OF COCHRAN'S THEOREM

TECHNICAL REPORT NO. 2

GEORGE P.J. STYAN
McGILL UNIVERSITY

SEPTEMBER 1982

U. S. ARMY RESEARCH OFFICE

CONTRACT DAAG29-82-K-0156

Also issued as Technical Report No. 56 under
Office of Naval Research Contract No. N00014-75-C-0442,
Department of Statistics, Stanford University.

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

A REVIEW AND SOME EXTENSIONS OF TAKEMURA'S GENERALIZATIONS OF COCHRAN'S THEOREM

by

George P.H. Styan
McGill University

This is a review of the Stanford Technical Report No.44 by Akimichi Takemura "On Generalizations of Cochran's Theorem and Projection Matrices" (August 1980). A number of extensions are also presented. In Section 1 we review the notions of direct sum of matrices and of vector spaces, and show the equivalence of several different definitions. We also examine Takemura's "independent" vector spaces and find that they coincide with the "disjoint" vector spaces considered by Rao and Yanai (1979). The connection of this concept with rank additivity is explored in Section 2. In Sections 3 and 4 we elaborate and extend several results of Takemura's on rank additivity and r-potent matrices and matrix polynomials. These results build on those given by Anderson and Styan (1980).

1. Direct Sums, Virtual Disjointness and Independence.

Let the vector spaces $\mathcal{U}_1, \ldots, \mathcal{U}_k$ be subspaces of a vector space $\mathcal{X}$ of $m\times 1$ column vectors and let the matrices $A_1, \ldots, A_k$ be defined by

$$(1.1) \qquad C(A_i) = \mathcal{U}_i, \qquad i = 1, \ldots, k,$$

where $C(.)$ denotes column space (range). The matrices $A_1, \ldots, A_k$ all have the same number of rows: m. Let A_i be $m\times n_i$ and write $n = \Sigma_1^k n_i$. Let

$$D = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_k \end{pmatrix} \tag{1.2}$$

be the $km \times n$ block-diagonal matrix of the A_i's.

The matrix D is defined as the *direct sum of the matrices* $A_1, \ldots, A_k$ (cf. Marcus and Minc, 1964, pages 5-6). This definition is also valid when the A_i do not all have the same number of rows. We may write

$$D = A_1 \oplus \ldots \oplus A_k. \tag{1.3}$$

Let the $km \times m$ partitioned matrix

$$K = (I_m, \ldots, I_m)', \tag{1.4}$$

cf. Anderson and Styan (1980, p.11). Then the $m \times n$ partitioned matrix

$$B = (A_1, \ldots, A_k) = K'D \tag{1.5}$$

spans the union of the vector spaces $\mathcal{U}_1, \ldots, \mathcal{U}_k$,

$$\mathcal{U} = \mathcal{U}_1 \cup \ldots \cup \mathcal{U}_k = C(A_1, \ldots, A_k) = C(B). \tag{1.6}$$

We define the matrices $A_1, \ldots, A_k$ and the vector spaces $\mathcal{U}_1, \ldots, \mathcal{U}_k$ to be *mutually virtually disjoint* whenever

$$r(A_1, \ldots, A_k) = \Sigma_1^k r(A_i), \tag{1.7}$$

where $r(.)$ denotes rank, cf. Styan (1981). We may write (1.7) as

$$r(K'D) = r(D) \tag{1.8}$$

or equivalently as

$$\dim(\mathcal{U}_1 \cup \ldots \cup \mathcal{U}_k) = \Sigma_1^k \dim\mathcal{U}_i, \tag{1.9}$$

where $\dim(.)$ denotes dimension of the vector space.

When the vector spaces $\mathcal{U}_1, \ldots, \mathcal{U}_k$ are *mutually virtually disjoint* then we define their union as the *direct sum of the vector spaces* $\mathcal{U}_1, \ldots, \mathcal{U}_k$, and we write

$$\mathcal{U}_1 \oplus \ldots \oplus \mathcal{U}_k = \mathcal{U}_1 \cup \ldots \cup \mathcal{U}_k = \mathcal{C}(A_1, \ldots, A_k) = \mathcal{C}(K'D). \tag{1.10}$$

Rao and Mitra (1971, page 3, lines 5-6) define the two vector spaces $\mathcal{U}_i$ and $\mathcal{U}_j$ to be *virtually disjoint* whenever their intersection

$$\mathcal{U}_i \cap \mathcal{U}_j = \{0\}, \tag{1.11}$$

the null vector only, or equivalently

$$r(A_i, A_j) = r(A_i) + r(A_j), \tag{1.12}$$

cf. Marsaglia and Styan (1974, page 272 (2.19)). Clearly (1.7) $\Rightarrow$ (1.12) for all $i \neq j$ and so (1.9) $\Rightarrow$ (1.11) for all $i \neq j$. The converse, however, holds in general only for $k = 2$; if $k = 3$ and

$$(1.13) \qquad A_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad A_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

then (1.12) holds for all $i \neq j$ but (1.7) does not.

Rao and Yanai (1979, page 2, section 2, lines 4-5) use (1.11) to define $\mathcal{U}_i$ and $\mathcal{U}_j$ "disjoint" and then (page 2, Definition 1) define $\mathcal{U}_1, \ldots, \mathcal{U}_k$ as "disjoint" whenever

$$(1.14) \qquad u \in \mathcal{U}_1 \cup \ldots \cup \mathcal{U}_k = C(A_1, \ldots, A_k) = C(B) = C(K'D)$$

has the unique representation

$$(1.15) \qquad u = \Sigma_1^k u_i; \qquad u_i \in \mathcal{U}_i, \ i = 1,\ldots,k.$$

Rao (1973, page 11 (vii)) uses this as a definition of "direct sum"; see also Takemura (1980, page 2, lines -6 and -7).

To see that (1.15) follows from (1.7) we write

$$(1.16) \qquad u = Bx^{(1)} = K'Dx^{(1)} = Bx^{(2)} = K'Dx^{(2)}.$$

Then

$$(1.17) \qquad K'Dx^{(1)} = K'Dx^{(2)} \quad \Rightarrow \quad Dx^{(1)} = Dx^{(2)},$$

which follows at once from (1.8) and the left-hand rank cancellation rule (Anderson and Styan (1980, page 12, Lemma 2.2)).

A connection between the definitions of matrix direct sum and vector-space direct sum is provided by the equality of the row spaces of $1\times n$ row vectors

$$(1.18) \qquad \mathcal{R}(K'D) = \mathcal{R}(D)$$

or equivalently

$$(1.19) \qquad \mathcal{R}(A_1, \ldots, A_k) = \mathcal{R}\begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_k \end{pmatrix} = \mathcal{R}(A_1 \oplus \ldots \oplus A_k).$$

This follows from (1.8) since $\mathcal{R}(K'D) \subseteq \mathcal{R}(D)$ always holds.

There seems to be no direct connection, however, between the vector-space direct sum $\mathcal{U}_1 \oplus \ldots \oplus \mathcal{U}_k = \mathcal{C}(A_1, \ldots, A_k)$ of $m\times 1$ vectors and the column space of $km\times 1$ vectors of the matrix direct sum $\mathcal{C}(A_1 \oplus \ldots \oplus A_k)$.

The "definition" given by Takemura (1980, page 2, paragraph 3, line 3) of $\mathcal{U}_1, \ldots, \mathcal{U}_k$ being <u>(linearly) independent</u> is that if

$$(1.20) \qquad u_i \in \mathcal{U}_i,\ i = 1,\ldots,k \quad \text{and} \quad \Sigma_1^k u_i = 0 \quad \text{then} \quad u_i = 0,\ i = 1,\ldots,k.$$

This is given by Rao and Yanai (1979, page 4) as the "result" (1): that if $A_1, \ldots, A_k$ are "disjoint" then

$$(1.21) \qquad \Sigma_1^k A_i x_i = 0 \iff A_i x_i = 0, \qquad i = 1,\ldots,k.$$

We may write (1.21) as

(1.22) $$K'Dx = 0 \iff Dx = 0,$$

which follows at once from (1.8) and the left-hand rank cancellation rule (cf. Anderson and Styan, 1980, page 12, Lemma 2.2).

Rao and Yanai (1979, page 5, Theorem 1) prove that if

(1.23) $$Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_k \end{pmatrix} = B^- = (A_1, \dots, A_k)^-$$

is a generalized inverse of B then

(1.24) $$A_i Y_i A_i = A_i,\ i = 1,\dots,k, \quad \text{and} \quad A_i Y_i A_j = 0 \ \text{ for all } \ i \neq j,$$

follow if and only if the A_i are "disjoint". We may write (1.23) as

(1.25) $$K'DYB = K'D = B$$

and (1.24) as

(1.26) $$DYB = D.$$

When the A_i are disjoint, (1.8) holds and so (1.25) $\Rightarrow$ (1.26) using the left-hand rank cancellation rule (Anderson and Styan, 1980, page 12, Lemma 2.2). Conversely (1.26) implies that $r(D) \le r(B) = r(K'D) \le r(D)$ and so (1.8) holds and the A_i are "disjoint".

This result from Rao and Yanai (1979, page 5, Theorem 1) closely parallels the result of Takemura (1980, page 4, Proposition 2.3): that if $P = P^2$ and $C(P) = \mathcal{U} = \mathcal{U}_1 \oplus \ldots \oplus \mathcal{U}_k$, with $\mathcal{U}_1, \ldots, \mathcal{U}_k$ "independent" then $P = \Sigma_1^k P_i$, with

$$(1.27) \qquad P_i^2 = P_i, \quad i = 1,\ldots,k, \text{ and } \quad P_i P_j = 0 \quad \text{for all } \; i \neq j,$$

and the P_i are unique. The correspondence between Takemura and Rao & Yanai is found by setting $P_i = A_i Y_i$; the assertion of uniqueness was not given by Rao & Yanai.

To prove Takemura's Proposition 2.3 let us write

$$(1.28) \qquad P = BY = \Sigma_1^k A_i Y_i$$

for some generalized inverse $Y = B^-$, cf. (1.23), and Theorem 3 of Marsaglia and Styan (1974, page 273). Setting $P_i = A_i Y_i$ then shows that (1.24) $\Rightarrow$ (1.27). The converse follows from $r(A_i) = \dim \mathcal{U}_i = r(P_i) = r(A_i Y_i)$ and the right-hand rank cancellation rule Lemma 2.1 in Anderson and Styan (1980, page 12).

To see that the P_i are unique we write

$$(1.29) \qquad P = BZ, \qquad \text{where} \quad Z = B^-$$

not necessarily equal to Y. We may do this in view of Theorem 3 of Marsaglia and Styan (1974, page 273). When the A_i are "independent" then

$$(1.30) \qquad P = K'DY = K'DZ \quad \Rightarrow \quad DY = DZ$$

using (1.8) and the left-hand rank cancellation rule, and so

$$(1.31) \qquad DY = \begin{pmatrix} A_1Y_1 \\ \vdots \\ A_kY_k \end{pmatrix} = \begin{pmatrix} P_1 \\ \vdots \\ P_k \end{pmatrix}$$

is uniquely determined.

2. Projection and Idempotent Matrices.

The square matrix A is idempotent whenever $A^2 = A$; such a matrix is called a "projection matrix" by Takemura (1980, page 2). An interesting discussion of this terminology is given by Ben-Israel and Greville (1974, page 51, footnote).

Takemura's Proposition 2.1 closely parallels Corollary 11.2 of Marsaglia and Styan (1974, page 283). Since the null space

$$(2.1) \qquad N(I - A) = \{x : (I-A)x = 0\} = \{x : Ax = x\}$$

it would seem more appropriate to replace (iii) and (iv), respectively, by

$$\text{(iii)}' \quad C(A) = N(I-A), \qquad \text{and} \qquad \text{(iv)}' \quad N(A) = C(I-A).$$

From (1.7) we see that

$$\text{(v)} \qquad C(A) \text{ and } C(I-A) \text{ are "independent"} \Longleftrightarrow r(A, I-A) = r(A) + r(I-A). \quad \ldots\text{(v)}'$$

Then it is obvious that (i) $\Leftrightarrow$ (ii) and that (iii)' $\Leftrightarrow$ (iv)', while (i) $\Leftrightarrow$ (v)' is (5.15) $\Leftrightarrow$ (5.21) in Corollary 11.2 of Marsaglia and Styan (1974, page 283), and (i) $\Leftrightarrow$ (iii)' is proved by Anderson and Styan (1980, pages 7-8, section 2.1).

Takemura's Proposition 2.2 (page 4) extends Theorem 1.1 of Anderson and Styan (1980, page 5). Let $A_1, \ldots, A_k$ be square matrices, not necessarily symmetric, and let $A = \Sigma_1^k A_i$. Consider the following statements

(a) $A_i^2 = A_i$, $i = 1,\ldots,k$,

(b) $A_i A_j = 0$ for all $i \neq j$,

(c) $A^2 = A$,

(d) $\Sigma_1^k r(A_i) = r(A)$,

(d1) $r(A_1, \ldots, A_k) = \Sigma_1^k r(A_i)$,

(d2) $C(A_1, \ldots, A_k) = C(A)$.

Then Takemura's Proposition 2.2 is that

(a), (b) $\Rightarrow$ (c), (d1), (d2).

In their (1.4), Anderson and Styan (1980, pages 5 and 11A) proved that

(a), (b) $\Rightarrow$ (c), (d).

That (d) $\Rightarrow$ (d1) is (6.1) $\Rightarrow$ (6.2) in Theorem 13 of Marsaglia and Styan (1974, pages 284-285). That (d1) $\Rightarrow$ (d2) follows at once since $C(A_1, \ldots, A_k)$ always contains $C(A)$, or equivalently $C(K'D) \supset C(K'DK)$.

Since (a), (b) $\Rightarrow$ (d1) Takemura (1980, page 5, lines -8 to -10) suggests that "it may well be justified to call projection matrices A_i, $i = 1,\ldots k$,

<u>(linearly) independent</u> if $A_i A_j = 0$ for all $i \neq j$." I think that this is not very wise since

$$\text{(a), (d1)} \not\Rightarrow \text{(b)};$$

i.e.,"independent" projections $A_1, \ldots, A_k$ are not necessarily "pairwise orthogonal": $A_i A_j = 0$ for all $i \neq j$. For example with $k = 2$, let

$$(2.2) \qquad A_1 = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

Takemura's Proposition 2.3 (page 4) has been discussed already in this review (see page 7).

Takemura's Proposition 2.4 (page 6) is well known and can easily be proved without diagonalizability (I prefer using "diagonability" and will do so below), cf. e.g., Rao (1973, page 28 (i)).

3. Cochran-type Theorems.

Takemura's Lemma 3.1 (page 7) builds on his Proposition 2.2 to show that

$$(d) \Longleftrightarrow (d1),\ (d2).$$

We have already shown that (d) $\Rightarrow$ (d1), (d2), cf. page 9 of this review. To go the other way let us write:

$$\begin{aligned} (d1) &: \quad r(K'D) = r(D) \\ (d2) &: \quad \mathcal{C}(K'D) = \mathcal{C}(K'DK) \\ (d) &: \quad r(K'DK) = r(D). \end{aligned}$$

Clearly (d2) $\Rightarrow r(K'D) = r(K'DK)$ and so (d1), (d2) $\Rightarrow$ (d).

Takemura's Theorem 3.2 (page 9) builds on Theorem 3.3 of Anderson and Styan (1980, page 24), but does not seem to be a "simpler (but equivalent) version". Let us write

$$
\begin{array}{llll}
(a)_r & : & A_i^r = A_i, & i = 1,\ldots,k, \\
(b) & : & A_iA_j = 0 & \text{for all } i \neq j, \\
(c)_r & : & A^r = A, & \\
(d) & : & \Sigma_1^k r(A_i) = r(A), & \\
(e2) & : & AA_i = A_iA, & i = 1,\ldots,k.
\end{array}
$$

The condition (e2) was used by Anderson and Styan (1980, page 18, Theorem 3.1). Takemura's Theorem 3.2 then shows that

$$(a)_r,\ (b) \Leftrightarrow (c)_r,\ (d),\ (e2).$$

That

$$(a)_r,\ (b) \Rightarrow (c)_r,\ (d)$$

was proved by Anderson and Styan (1980, page 18, Theorem 3.1), while

$$(b) \Rightarrow (e2)$$

always holds. To go the other way, we note first that

$$(d),\ (e2) \Rightarrow (b).$$

This was proved for symmetric matrices by Luther (1965, page 684, Theorem 1) and for square matrices not necessarily symmetric by Marsaglia (1967, page 461, Theorem 3); see also Marsaglia and Styan (1974, page 286, Theorem 15(II)). Hence

$$(c)_r, (d), (e2) \Rightarrow (b), (c)_r \Rightarrow (a)_r, (b)$$

using (3.23) on page 25 of Anderson and Styan (1980). In their Theorem 3.3 Anderson and Styan (1980, page 24) proved that

$$(a)_r, (b) \Leftrightarrow (c)_r, (d), (e2)_r,$$

where

$$(e2)_r \quad : \quad A^{r-2}A_i = A_iA^{r-2}, \qquad i = 1,\ldots,k.$$

Since

$$(e2) \Rightarrow (e2)_r \qquad \text{and} \qquad (b) \Rightarrow (e2)$$

always hold, and since

$$(d), (e2) \Rightarrow (b)$$

(see page 12 above) it seems to me that Takemura's Theorem 3.2 is equivalent to the Anderson and Styan Theorem 3.3.

4. Matrix Polynomials.

Takemura's Theorem 3.3 (page 10) extends the Anderson and Styan Theorem 3.3 by replacing their (a) and (c) conditions by a matrix polynomial P(.) of degree at least 2. Let us write

$$
\begin{array}{lcll}
(a)_{P0} & : & P(A_i) = 0, & i = 1,\dots,k, \\
(a)_{P1} & : & A_i P(A_i) = 0, & i = 1,\dots,k, \\
(b) & : & A_i A_j = 0 & \text{for all } i \neq j, \\
(c)_P & : & P(A) = 0, & \\
(d) & : & \Sigma_1^k r(A_i) = r(A), & \\
(e2) & : & AA_i = A_i A, & i = 1,\dots,k.
\end{array}
$$

Takemura's Theorem 3.3 then shows that

$$(c)_P,\ (d),\ (e2) \Rightarrow (a)_P,\ (b),$$

where

$$(a)_P = \begin{cases} (a)_{P0} & \text{if} \quad P(0) = 0, \\ (a)_{P1} & \text{if} \quad P(0) \neq 0. \end{cases}$$

Since $(d), (e2) \Rightarrow (b)$ always holds, cf. page 12 above, Takemura's Theorem 3.3 reduces to

$$(b),\ (c)_P,\ (d) \Rightarrow (a)_P. \tag{4.1}$$

We strengthen (4.1) by showing that

$$(4.2) \qquad (b),\ (c)_P,\ (d) \Rightarrow (a)_P^* ,$$

where

$$(a)_P^* \ : \quad P(A_i) = c(I - A^- A_i), \qquad i = 1,\dots,k,$$

$P(0) = cI$, and A^- is a generalized inverse of A. Then $(a)_P^* \Rightarrow (a)_{P0}$ if $c = 0$. To prove (4.2) we note that when (b) holds

$$(4.3) \qquad A^p = \Sigma_1^k A_i^p$$

for any positive integer p, and so

$$(4.4) \qquad P(A) = cI + \Sigma_1^k [P(A_i) - cI].$$

When (d) holds we may apply Theorem 1.2 of Anderson and Styan (1980, page 6) to find

$$(4.5) \qquad A_i A^- A_i = A_i, \quad i = 1,\dots,k, \text{ and } A_i A^- A_j = 0 \quad \text{for all } i \neq j.$$

Postmultiplying (4.4) by $A^- A_i$ and applying (4.5) yields

$$(4.6) \qquad [P(A)]A^- A_i = cA^- A_i + [P(A_i) - cI], \qquad i = 1,\dots,k,$$

and so $(c)_P \Rightarrow (a)_P^*$. Premultiplying $(a)_P^*$ by A_i yields $(a)_{P1}$.

Takemura's Lemma 4.1 (page 14) shows that if $P(x) = 0$ has no multiple root and if $(c)_P$: $P(A) = 0$ holds, then A is diagonable. Since (by definition) the minimal polynomial must divide $P(A)$ it follows that when $P(A)$ has no multiple root then neither does the minimal polynomial $\Phi(A)$. This characterizes diagonability of the matrix A, cf. Mirsky (1955, page 297, Theorem 10.2.5), and so the condition in Lemma 4.1 does not seem to be as strong as Takemura claims on page 13 (line -1).

Takemura's Theorem 4.1 is not new. It contains his Lemma 4.1 and the decomposition into "independent projections", more usually known as "principal idempotents" (cf. Ben-Israel and Greville, 1974, pages 52-55, especially Theorem 9 and Exercise 27).

To Takemura's Corollary 4.1 should be added $\Sigma_1^r H_j = I$.

Takemura's Theorem 4.2 (pages 15-16) considers a polynomial $P(x)$ with no multiple root and then extends his Theorem 3.3 to show that

$$(c)_P,\ (d),\ (e2) \Rightarrow A_i = \sum_{j=1}^{\ell} \lambda_j H_{ij},$$

where $\lambda_1, \ldots, \lambda_\ell$ are the nonzero characteristic roots of A and the H_{ij} are "independent projections" (principal idempotents). In fact $H_{ij}H_{i'j'} = 0$ whenever $i \neq i'$ and/or $j \neq j'$. The condition that $P(x)$ has no multiple root assures that $(c)_P \Rightarrow A$ diagonable. Since (d), (e2) $\Rightarrow$ (b) we see the interesting result (essentially Takemura's Theorem 4.4, page 21) that

$$\text{(4.7)} \qquad A \text{ diagonable, (b), (d)} \Rightarrow A_i \text{ diagonable, } i = 1,\ldots,k.$$

[We must still include (d) in (4.7) since (b) $\not\Rightarrow$ (d) in general; however (b) always implies (e2).]

We may extend (4.7) by noting that a matrix A is diagonable if and only if all its characteristic roots are regular (cf. Mirsky, 1955, page 294, Theorem 10.2.3). The characteristic root λ of the matrix A is said to be regular whenever its geometric and algebraic multiplicities are equal (cf. Mirsky, 1955, page 294, Definition 10.2.1). The algebraic multiplicity of λ is the multiplicity of λ as a root of the characteristic equation; the geometric multiplicity is the nullity of the matrix $A - \lambda I$.

Let A be $n \times n$ of rank r, and let A_i have rank r_i, $i = 1,\ldots,k$. Let $\lambda_1, \ldots, \lambda_\ell$ be the nonzero characteristic roots of A. Let m_{ij} be the algebraic multiplicity and g_{ij} the geometric multiplicity of λ_j as a characteristic root of A_i, so that, cf. Mirsky (1955, page 214, Theorem 7.6.1),

$$n \geq m_{ij} \geq g_{ij} \geq 0; \qquad i = 1,\ldots,k, \qquad j = 1,\ldots,\ell. \tag{4.8}$$

Let m_{0j} be the algebraic multiplicity and g_{0j} the geometric multiplicity of λ_j as a characteristic root of A. Then

$$n \geq m_{0j} \geq g_{0j} \geq 1; \qquad j = 1,\ldots,\ell. \tag{4.9}$$

Let m_{i0} be the algebraic multiplicity and g_{i0} the geometric multiplicity of 0 as a characteristic root of A_i, $i = 1,\ldots,k$. Then

$$m_{i0} = n - m_{i\cdot} = n - \sum_{j=1}^{\ell} m_{ij}, \qquad i = 1,\ldots,k, \tag{4.10}$$

$$g_{i0} = n - r_i = n - \mathrm{rank}(A_i), \qquad i = 1,\ldots,k. \tag{4.11}$$

Hence

(4.12) $$n \geq r_i \geq m_{i\cdot} \geq 0; \qquad i = 1,\ldots,k.$$

Let m_{00} be the algebraic multiplicity and let g_{00} be the geometric multiplicity of 0 as a characteristic root of A. Then

(4.13) $$m_{00} = n - m_{0\cdot} = n - \sum_{j=1}^{\ell} m_{0j} ,$$

(4.14) $$g_{00} = n - r = n - \mathrm{rank}(A).$$

Hence

(4.15) $$n \geq r \geq m_{0\cdot} \geq 0.$$

Let

(4.16) $$A_i = B_i C_i', \qquad i = 1,\ldots,k,$$

be full rank decompositions, so that B_i and C_i are both $n \times r_i$ of rank r_i.
Then

(4.17) $$A = \Sigma_1^k A_i = \Sigma_1^k B_i C_i' = BC',$$

where

(4.18) $$B = (B_1, \ldots, B_k) \quad \text{and} \quad C = (C_1, \ldots, C_k)$$

are both $n \times \Sigma_1^k r_i$.

Now suppose that

$$\text{(b)} \qquad A_i A_j = 0 \qquad \text{for all} \quad i \neq j$$

holds. Then $C_i'B_j = 0$ for all $i \neq j$ and so

$$(4.19) \qquad C'B = \begin{pmatrix} C_1'B_1 & & \\ & \ddots & \\ & & C_k'B_k \end{pmatrix} = C_1'B_1 \oplus \cdots \oplus C_k'B_k,$$

the direct sum of the $C_i'B_i$.

Let $am_j(A)$ denote the algebraic multiplicity and let $gm_j(A)$ denote the geometric multiplicity of λ_j as a characteristic root of A. Then using the fact that the matrices FG and GF have the same nonzero characteristic roots (cf. Mirsky, 1955, page 200, Theorem 7.2.3), we may write for $j = 1,\ldots,\ell$

$$(4.20) \qquad m_{0j} = am_j(A) = am_j(BC') = am_j(C'B) = \sum_{i=1}^{k} am_j(C_i'B_i)$$

$$= \sum_{i=1}^{k} am_j(B_iC_i') = \sum_{i=1}^{k} m_{ij},$$

while

$$(4.21) \qquad m_{0\cdot} = \sum_{j=1}^{\ell} m_{0j} = \sum_{i=1}^{k} \sum_{j=1}^{\ell} m_{ij} = \sum_{i=1}^{k} m_{i\cdot}$$

so that

$$(4.22) \qquad n - am_0(A) = \Sigma_1^k [n - am_0(A_i)].$$

We now use the result that the matrices $FG - I$ and $GF - I$ have the same nullity, cf. Ouellette (1981, equation (4.147)). Let $\nu(.)$ denote nullity. Then for each $j = 1,\ldots,\ell$

$$\begin{aligned} (4.23)\qquad g_{0j} &= gm_j(A) = \nu(A - \lambda_j I) = \nu(BC' - \lambda_j I) = \nu(C'B - \lambda_j I) \\ &= \sum_{i=1}^{k} \nu(C_i'B_i - \lambda_j I_{r_i}) = \sum_{i=1}^{k} \nu(B_iC_i' - \lambda_j I_n) \\ &= \sum_{i=1}^{k} \nu(A_i - \lambda_j I_n) = \sum_{i=1}^{k} g_{ij}. \end{aligned}$$

Hence when (b) holds all the nonzero characteristic roots of the A_i, $i = 1,\ldots,k$, must be characteristic roots of A, and all the nonzero characteristic roots of A must be characteristic roots of A_i for some i.

Furthermore, since $g_{ij} \le m_{ij}$ from (4.8) we obtain

$$(4.24)\qquad g_{0j} = \sum_{i=1}^{k} g_{ij} \le \sum_{i=1}^{k} m_{ij} = m_{0j}; \qquad j = 1,\ldots,\ell,$$

and so for each $j = 1,\ldots,\ell$

$$(4.25)\qquad g_{0j} = m_{0j} \Rightarrow g_{ij} = m_{ij}; \qquad i = 1,\ldots,k.$$

Thus if λ_j is a regular nonzero characteristic root of A then when (b) holds λ_j is also a regular characteristic root† of each A_i, $i = 1,\ldots,k$. This does not, however, hold true for the 0 characteristic root of A, for if $k=2$ and

† Notice that $g_{ij} = 0 \Leftrightarrow m_{ij} = 0$; we will then speak of λ_j as a regular characteristic root of A_i even though A_i does not have λ_j as a root.

$$(4.26) \qquad A_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad A_2 = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}$$

then (b) holds and $A = 0$ so 0 is a regular characteristic root of A, but 0 is clearly not a regular characteristic root of either A_1 or of A_2. The equation (4.23) is no longer valid when $\lambda_j = 0$ since the matrices FG and GF do not necessarily have the same nullity (or rank). However

$$(4.27) \qquad n - gm_0(A) = r \leq \Sigma_1^k r_i = \Sigma_1^k [n - gm_0(A_i)].$$

Now suppose that in addition

$$(d) \qquad \Sigma_1^k r(A_i) = r(A)$$

holds. Substitution in (4.27) then yields

$$(4.28) \qquad n - gm_0(A) = \Sigma_1^k [n - gm_0(A_i)]$$

and so

$$(4.29) \qquad \begin{aligned} gm_0(A) &= n - \Sigma_1^k [n - gm_0(A_i)] \\ &\leq n - \Sigma_1^k [n - am_0(A_i)] = n - \Sigma_1^k m_{i\cdot} = n - m_{0\cdot}. \end{aligned}$$

Hence if 0 is a regular characteristic root of A then 0 must be a regular characteristic root of A_i, for all $i = 1,\dots,k$. We have, therefore, proved the following stronger version of Takemura's Theorems 4.2 and 4.4 and our (4.7).

THEOREM 4.1. *Let* $A_1, \ldots, A_k$ *be* $n \times n$ *matrices, not necessarily symmetric, and let* $A = \Sigma_1^k A_i$. *Suppose that*

(b) $$A_i A_j = 0 \quad \textit{for all} \quad i \neq j.$$

Then the set of nonzero characteristic roots of A *coincides with the set of all the nonzero characteristic roots of all the* A_i, $i = 1,\ldots,k$. *Furthermore, the nonzero characteristic root* λ *of* A *is regular if and only if* λ *is a regular nonzero characteristic root of each* A_i, $i = 1,\ldots,k$.

If, in addition,

(d) $$\Sigma_1^k \mathrm{rank}(A_i) = \mathrm{rank}(A),$$

then the characteristic root 0 *of* A *is regular if and only if* 0 *is a regular characteristic root of each* A_i, $i = 1,\ldots,k$. *Equivalently,*

(4.30) $$\mathrm{rank}(A^2) = \mathrm{rank}(A) \iff \mathrm{rank}(A_i^2) = \mathrm{rank}(A_i), \quad i = 1,\ldots,k.$$

Mäkeläinen and Styan (1976, Lemma 2) have shown that the zero characteristic root of a matrix A is regular if and only if $r(A^2) = r(A)$. Such a matrix A is said to have index 1, cf. Ben-Israel and Greville (1974, page 169). A direct proof of (4.30) goes as follows. Let

(4.31) $$D = A_1 \oplus \ldots \oplus A_k \quad \text{and} \quad K = (I, \ldots, I)',$$

cf. (1.3) and (1.4) on page 2. Then under (b) $DKK'D = D^2$ so that $A^2 = K'D^2K$, while (d) yields

(4.32) $$r(D) = r(K'DK) = r(K'D) = r(DK)$$

using Sylvester's Law of Nullity, cf. Anderson and Styan (1980, page 11A). Thus under (b) and (d) the result (4.30) becomes

(4.33) $$r(K'D^2K) = r(D)$$

if and only if

(4.34) $$r(D^2) = r(D).$$

That (4.33) $\Rightarrow$ (4.34) follows at once from $r(K'D^2K) \leq r(D^2) \leq r(D)$. To go the other way suppose that (4.34) holds. Then using (4.32) and the rank cancellation rules (2.10) and (2.11) in Theorem 2 of Marsaglia and Styan (1974, page 271), we obtain $r(K'D^2K) = r(D^2K) = r(D^2)$, and so (4.34) $\Rightarrow$ (4.33).

We see also that if (b) and (d) hold then for any positive integer p (4.32) implies that $r(K'D^pK) = r(D^p)$ and so we have proved that

THEOREM 4.2. *Let the* $n \times n$ *matrices* $A_1, \ldots, A_k$ *and* A *be defined as in Theorem 4.1 and suppose that both* (b) *and* (d) *hold. Then for any positive integer* p

(4.35) $$\Sigma_1^k \operatorname{rank}(A_i^p) = \operatorname{rank}\left(\Sigma_1^k A_i^p\right).$$

Takemura's Theorem 4.3 (page 16) is well known; for a nice proof see Mirsky (1955, page 318, Theorem 10.6.3).

We may strengthen Takemura's Corollary 4.2 (page 20) by allowing the polynomial $P(x)$ to possibly have multiple roots and by not requiring that $P(0) = 0$. Instead then let $P(0) = cI$ as on page 15 of this review. Thus from (4.3) we find that if $P(A_i) = 0$, $i = 1,\ldots,k$, then under (b) $\equiv$ (i)

$$P(A) = -(k-1)cI. \tag{4.36}$$

Since (b) always implies (e2) $\equiv$ (iv), it remains to consider when does (b) imply (d) $\equiv$ (v). That (b) $\not\Rightarrow$ (d) in general has already been noted (top of page 17 of this review; see also Anderson and Styan (1980, page 5)). Takemura's Corollary 4.2 shows that

$$\text{(b) and } A_i \text{ diagonable, } i = 1,\ldots,k \Rightarrow \text{(d)}. \tag{4.37}$$

This result, however, is implied by

$$\text{(b) and } \operatorname{rank}(A_i^2) = \operatorname{rank}(A_i), \quad i = 1,\ldots,k \Rightarrow \text{(d)}, \tag{4.38}$$

which was proved by Marsaglia and Styan (1974, page 286, Theorem 15(I)). We recall that $\operatorname{rank}(A_i^2) = \operatorname{rank}(A_i)$ means that the zero characteristic root of A_i is regular, while A_i diagonable means that all the characteristic roots are regular, cf. pages 17 and 22 of this review.

ACKNOWLEDGEMENTS

This research was supported in part by the Natural Sciences and Engineering Research Council Canada, Grant Number A7274, and by the Gouvernement du Québec, Programme de formation de chercheurs et d'action concertée, subvention no. EQ-961.

REFERENCES

Anderson, T.W., and Styan, George P.H. (1980). Cochran's theorem, rank additivity, and tripotent matrices. Technical Report No.44, Dept. of Statistics, Stanford University. Revised version published in *Statistics and Probability: Essays in Honor of C.R. Rao* (G. Kallianpur, P.R. Krishnaiah, & J.K. Ghosh, *eds.*), North-Holland, Amsterdam, 1982, pp. 1-23.

Ben-Israel, Adi, and Greville, Thomas N.E. (1974). *Generalized Inverses: Theory and Applications*. Wiley, New York. Reprinted with corrections, Robert E. Krieger, Huntington, New York, 1980.

Luther, Norman Y. (1965). Decomposition of symmetric matrices and distributions of quadratic forms. *The Annals of Mathematical Statistics*, 36, 683-690.

Mäkeläinen, Timo, and Styan, George P.H. (1976). A decomposition of an idempotent matrix where nonnegativity implies idempotence and none of the matrices need be symmetric. *Sankhyā Series A*, 38, 400-403.

Marcus, Marvin, and Minc, Henryk (1964). *A Survey of Matrix Theory and Matrix Inequalities*. Allyn and Bacon, and Prindle, Weber and Schmidt, Boston.

Marsaglia, George (1967). Bounds on the rank of the sum of matrices. *Transactions of the Fourth Prague Conference on Information Theory, Statistical Decision Functions, and Random Processes* (Prague, August 31 - September 11, 1965), Czechoslovak Academy of Sciences, 455-462.

Marsaglia, George, and Styan, George P.H. (1974). Equalities and inequalities for ranks of matrices. *Linear and Multilinear Algebra*, 2, 269-292.

Mirsky, L. (1955). *An Introduction to Linear Algebra*. Oxford University Press.

Ouellette, Diane Valérie (1981). Schur complements and statistics. *Linear Algebra and its Applications (Letters in Linear Algebra)*, 36, 187-295. [M.Sc. thesis, Dept. of Mathematics, McGill University, Montréal, July 1978.]

Rao, C. Radhakrishna (1973). *Linear Statistical Inference and Its Applications*. Second Edition. Wiley, New York.

Rao, C. Radhakrishna, and Mitra, Sujit Kumar (1971). *Generalized Inverse of Matrices and Its Applications*. Wiley, New York.

Rao, C. Radhakrishna, and Yanai, Haruo (1979). General definition and decomposition of projectors and some applications to statistical problems. *Journal of Statistical Planning and Inference*, 3, 1-17.

Styan, George P.H. (1981). Review of Rao and Yanai (1979). *Mathematical Reviews*, in press.

Takemura, Akimichi (1980). On generalizations of Cochran's theorem and projection matrices. Technical Report No.44, Dept. of Statistics, Stanford University.

GPHS : 01.01.81

Updated: 09.07.82

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Date Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 2	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A Review and Some Extensions of Takemura's Generalizations of Cochran's Theorem		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) George P.H. Styan		8. CONTRACT OR GRANT NUMBER(s) DAAG29-82-K-0156
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics - Sequoia Hall Stanford University Stanford, California 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS P-19065-M
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office Post Office Box 12211 Research Triangle Park, NC 27709		12. REPORT DATE September 1982
		13. NUMBER OF PAGES 26
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

Also issued as Technical Report No. 56 under Office of Naval Research Contract No. N00014-75-C-0442.

18. SUPPLEMENTARY NOTES

THE VIEW, OPINIONS, AND/OR FINDINGS CONTAINED IN THIS REPORT ARE THOSE OF THE AUTHOR(S) AND SHOULD NOT BE CONSTRUED AS AN OFFICIAL DEPARTMENT OF THE ARMY POSITION, POLICY, OR DECISION, UNLESS SO DESIGNATED BY OTHER DOCUMENTATION.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Cochran's theorem, rank additivity, r-potent matrices, direct sum, projection matrices, diagonability, matrix polynomial.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Generalizations of Cochran's theorem by Takemura are discussed in connection with results by Anderson and Styan and other works by Rao and Yanai and Marsaglia and Styan. A number of further extensions are presented.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Date Entered)